

The Academic Impact of Enrollment in International Baccalaureate Diploma Programs: A Case Study of Chicago Public Schools

ANNA ROSEFSKY SAAVEDRA

RAND Corporation

Background: *In schools accredited as “IB World Schools” by the International Baccalaureate Organization (IBO), teachers use IB curriculum and pedagogy to teach a range of courses that are intended to prepare IB-enrolled students for college. Over the past 18 years, the number of U.S. schools that implement IB programs has increased nearly tenfold, from 133 in 1994 to 1,390 in 2013. Despite the IB program’s rapid expansion, little is known about whether IB enrollment causally improves students’ academic outcomes, including their high school academic achievement, probability of high school graduation and/or subsequent probability of college enrollment.*

Purpose: *This study examines whether enrollment in the IB Diploma Program increases students’ academic achievement as measured by their composite ACT college admissions examination scores, probability of high school graduation, and probability of college enrollment, and whether the estimates differ by gender.*

Setting, Population, & Data: *This study uses data on the demographic characteristics, IB enrollment status, ACT scores, high school graduation status and college enrollment status of 20,422 students attending 13 Chicago Public Schools (CPS) high schools from 2002-2008. Data sources include the CPS and the National Student Clearinghouse.*

Research Design: *The analytic strategy is to first assume students are selected into the IB Diploma Program based on their observed characteristics, and then to use a propensity score approach to estimate the impact of IB enrollment on three measures of students’ academic success. The second step, following Rosenbaum (2002), is to test the sensitivity of the estimates to different levels of selection bias.*

Results: *This study shows that IB enrollment increases students’ academic achievement, probability of high school graduation and probability of college enrollment. Though selection*

bias may contribute to overstating the propensity score estimates, the conclusion from the sensitivity analyses is that it is unlikely that this internal-validity challenge negates the principal finding. All estimates are greater for boys than for girls. Calculations demonstrate that the IB Diploma Program is a cost-effective way to increase high school graduation rates.

Conclusions: *The results are valuable for three reasons. First, they provide valuable information with which to make decisions about future investments in IB. Second, they contribute to knowledge of the means through which to improve high school education for disadvantaged urban youth. Finally, the results suggest that IB enrollment is especially beneficial for boys, for whom the probability of graduating from high school and enrolling in college—in CPS and at the national level—is substantially less than for girls.*

Originally founded as a private means for diplomats' children to earn an internationally recognized high school diploma, today the International Baccalaureate (IB) Diploma Program¹ is one of the fastest-growing curricular innovations in U.S. public-school districts. In schools accredited as "IB World Schools" by the IB Organization (IBO), teachers use IB curriculum and pedagogy to teach a range of courses that are intended to prepare IB-enrolled students for college. This study examines whether IB enrollment actually prepares students for college. It asks whether enrollment in the IB Diploma Program increases students' academic achievement as measured by their composite ACT college admissions examination scores,² probability of high school graduation, and probability of college enrollment, and whether the estimates differ by gender.

INTRODUCTION

THE INTERNATIONAL BACCALAUREATE DIPLOMA PROGRAM

Over the past 20 years, the number of U.S. schools that implement IB programs has increased nearly tenfold, from 133 in 1994 to 1,390 in 2013. By 2009, public schools offered over 90% of U.S. high school level IB programs (IBO, 2012). What began as a program that a few private schools offered to wealthy children today reaches a much broader student audience.

This growth in IB implementation is based predominantly on the program's perceived academic rigor and success as a college-preparatory intervention (Byrd, Ellington, Gross, Jago, & Stern, 2007; Cech, 2008; Mathews & Hill, 2005). This reputation, augmented through recent governmental support, suggests that the growth trend will continue. For example, through the reauthorization of the Elementary and Secondary Schools Act, the U.S. Department of Education has proposed to fund IB programs as part of its objective to "increase access to accelerated

learning opportunities” (U.S. Department of Education, 2011a, p. 29). State-level education departments are also promoting IB implementation as part of their “Race to the Top” strategies (e.g., FL, MA, ME, NH; U.S. Department of Education, 2011b).

The IB school-level accreditation process and subsequent requirements for students and teachers are rigorous (Byrd et al., 2007; Mathews & Hill, 2005; U.S. Department of Education, 2011a). To earn accreditation from the IBO, schools must demonstrate adherence to IBO’s curricular, pedagogical, mission-based, and ongoing professional-development requirements as stated in the current version of the “Handbook of Procedures for the Diploma Program” (IBO, 2013). Typically, the IB accreditation process requires extensive faculty and staff participation, takes several years to complete, and, if successful, results in authorized “IB World School” status. IB World-School teachers must use IB curricular and pedagogical materials to teach IB courses, and participate in ongoing professional development and self-review processes.

Students enroll in the official IB Diploma Program in 11th grade. Over the course of their final two years of high school, to be eligible to earn the IB Diploma, students must: enroll in six core IB courses and the IB Theory of Knowledge (TOK) epistemology course; participate in the weekly Creativity, Action and Service (CAS) requirement; and write a 4,000-word “extended essay.” They must also score above a defined threshold on IBO-created and administered examinations in six subjects including languages, social studies, experimental sciences, and mathematics. Approximately 10,000 IB-certified examiners, in more than 100 countries worldwide, mark the examinations (IBO, 2013).

Despite the IB program’s rapid expansion, little is known about whether enrollment in the IB Diploma Program improves students’ academic outcomes, including their high school academic achievement, probability of high school graduation, and/or subsequent probability of college enrollment (Foust, Hertberg-Davis, & Callahan, 2009; Jackson, 2010; Kyburg, Hertberg-Davis, & Callahan, 2007; Roderick, Nagaoka, Coca, & Moeller, 2009). To date, only one study has examined the relationship between IB enrollment and students’ high school achievement, graduation, and subsequent college entry (Roderick et al., 2009). The purpose of that study, however, was to examine whether Chicago Public Schools’ (CPS) students who enrolled in the IB Diploma Program, Advanced Placement, or other honors-level courses subsequently enrolled in colleges that were appropriately competitive. In contrast to the present study, according to Roderick et al., “[their] report is not intended to be a rigorous evaluation of selective enrollment schools, IB programs or AP initiatives” (2009, p. 2). Another study uses a propensity-score analysis methodology to

measure the impact of IB enrollment on students' college outcomes, including whether they enrolled in four-year colleges and/or selective four-year colleges, as well as the extent to which they persisted in four-year colleges (Coca, Kelley-Kemple, Roderick, & Nagaoka, 2011). In contrast to the present study, Coca et al. do not address the same outcomes of interest or use a rigorous sensitivity analysis to test the robustness of the propensity-score estimates.

In addition, though other studies have found differences in IB course-taking patterns by gender (Kieran, 2007; Stables & Sian, 1995) and that, compared to boys, the probability is greater that girls will complete rigorous or midlevel curricula (U.S. Department of Education, 2007), no studies have investigated whether the probability or impact of enrollment in IB differs by gender.

THE IB DIPLOMA PROGRAM IN CHICAGO PUBLIC SCHOOLS

This study examines whether overall and by gender, 11th-grade CPS students' IB enrollment improves their academic achievement, probability of high school graduation, and probability of college enrollment. CPS is an ideal setting in which to address this issue for three reasons. First, CPS, at the forefront of IB implementation nationwide, currently offers the IB Diploma Program in 14 high schools, more than twice that of any other U.S. school district.³ Next, the system retains historical records of its students' IB enrollment status, achievement test scores, high school graduation status, and college-enrollment information. Finally, since the average CPS student is low income and minority, evidence documenting the IB enrollment impact in this setting would enhance knowledge of ways to improve secondary education for disadvantaged urban youth.

In the early 1990s, levels of academic achievement in most CPS high schools were dismal. Ninth- and 11th-grade students' achievement scores revealed academic performance that was, on average, more than a year behind grade level when compared to national averages (Chicago School Reform Board of Trustees, 1999). The district's high school graduation rate in 1995 was less than 50% of an entering cohort, relative to the national average of 76% (Heckman & LaFontaine, 2010).⁴ Among Illinois students who took the 1994-1995 ACT college entrance examination—50% of CPS seniors and 57% statewide—CPS students lagged behind their peers statewide by nearly five points out of a maximum score of 36 (Rice, 1995).

Boys who attended CPS neighborhood schools fared particularly poorly. For example, in the population of CPS students who were 13 years old in 1998, only 39% of African-American, 51% of Latino, and 55% of White boys had graduated from high school by 2004, when this cohort

was 19 years old. In comparison, 57% of African-American girls, 65% of Latino girls, and 71% of White girls in the 1998 cohort graduated by 2004 (Allensworth, 2005). Arguably, the need for school reform in CPS was even greater for boys than for girls.

In 1995, CPS operated only three “selective enrollment” magnet schools, enrolling less than 5% of the CPS high school population. Incoming freshmen, of whom an estimated 40% applied to magnet schools, annually vied to be one of the fortunate 5% (Roderick et al., 2009). Though five new magnet schools opened between 1995 and 1999, they still represented only 6% of total CPS high school enrollment (Allensworth & Rosenkranz, 2000). Given the high demand and low supply of magnet school slots, annually thousands of motivated, academically talented CPS students could not access academically rigorous programs.

If parents had the financial means, they sent their children to private or parochial schools or moved to the suburbs. Students who left the CPS system to attend private, parochial, or suburban schools were disproportionately higher academic achievers. Whereas 27% of students who scored above the 60th percentile on state standardized tests left the CPS system in 1995 between seventh and ninth grades, only 12% of students who scored below the 60th percentile left (Allensworth & Rosenkranz, 2000). This phenomenon, known as “brain-drain,” left a concentration of economically disadvantaged and lower-performing students in the neighborhood schools.

From 1981 through 1997, Lincoln Park was the only public, non-magnet high school, out of more than 90 in CPS, to offer the IB Diploma Program. From its inception, Lincoln Park IB students’ academic achievement far surpassed that of students from other neighborhood high schools and even from most of the CPS selective-enrollment high schools (Roderick et al., 2009).

Driven by Lincoln Park’s IB success, beginning in 1997, CPS rapidly increased the number of IB-accredited programs offered in Chicago neighborhood high schools. The 12 IB schools that opened between 1997 and 2004 were evenly distributed across CPS’s six school catchment zones, such that with just one exception, each zone received two IB programs. This even catchment distribution made the IB option geographically accessible to students in all neighborhoods. Despite the even catchment distribution, IB schools are more concentrated in North and Southwest Chicago, which have high concentrations of Latino families and adults with little or no college education (Coca et al., 2011). CPS’s primary goal for widely implementing the IB Diploma Program was to increase predominantly low-income and minority students’ access to academically rigorous curricula in neighborhood high schools and,

thereby, to increase their academic achievement and college preparedness (Spittle, Leven, & Roderick, 2008).

The 13 CPS high school level IB programs included in this study serve primarily low-income minority students: CPS administrative data show that 65% of CPS students enrolled in IB are eligible to receive free or reduced-price lunch and 74% are racial minorities. Among non-IB (11th grade) students in the district, in comparison, 78% are eligible to receive free or reduced-price lunch and 85% are racial minorities. Therefore, both groups are majority-minority and low-income, though on average, CPS IB students are slightly more advantaged than non-IB students.

The IBO does not dictate how schools determine student eligibility for enrollment in IB programs. The CPS policy is to require all students who think they would like to enroll in the official 11th- through 12th-grade IB Diploma Program in their geographically determined neighborhood high school to apply in the spring of their eighth-grade year so that they can participate in an IB-preparation track during ninth and 10th grades. The nature of the preparation track differs by CPS high school. Several offer accredited IB-preparation programs (the IB Middle Years Program), while others offer AP courses or other honors classes.

As part of the IB application requirement for the cohorts included in this study, CPS required students to submit their seventh-grade Iowa Test of Basic Skills (ITBS) mathematics and reading scores. For the first few years of widespread IB Diploma Program implementation, the stated CPS policy was to admit only students scoring above the 60th percentile on both ITBS tests. According to the 1997-2008 CPS IB Diploma Program administrator, this policy evolved over time such that by several years post-widespread IB implementation, most schools began to permit IB enrollment among students scoring above the 50th percentiles on the reading and mathematics portions of the seventh-grade test in current use (personal correspondence, March 26, 2009).⁵

RESEARCH QUESTIONS

Currently, policy makers, administrators, teachers, parents, and students lack credible evidence on whether IB enrollment increases students' academic achievement, probability of high school graduation, and probability of college enrollment, and whether effects might differ by gender. Such evidence could inform decisions about future investments in the IB Diploma Program throughout Chicago, and in U.S. public schools.

This need for evidence motivates the study's research questions:

- 1a. Does CPS students' enrollment in the IB Diploma Program in 11th grade increase the probability that they will take the ACT college entrance examination?

- 1b. Does CPS students' enrollment in the IB Diploma Program in 11th grade improve their high school achievement, as measured by their ACT scores?
2. Does CPS students' enrollment in the IB Diploma Program in 11th grade increase the probability that they will graduate from high school?
3. Does CPS students' enrollment in the IB program in 11th grade increase the probability that they will enroll in a two- or four-year college at some point during the first two years after their graduation from high school?
4. Do potential effects of IB enrollment on students' probability of taking the ACT examination, academic achievement as measured by ACT scores, probability of graduation from high school, and probability of enrollment in college, respectively, differ by gender?

METHODS

DATASET

The data analyzed in this study contains information merged from two sources, from the CPS themselves and from the National Student Clearinghouse (NSC), a nonprofit organization that verifies students' enrollment in, and graduation from, 92% of U.S. colleges (National Student Clearinghouse, 2011).⁶ The CPS dataset contains records for the 20,422 students who attended one of the 13 CPS high schools that offered the IB Diploma Program to the 2002-2003, 2003-2004, 2004-2005, and 2006-2007 11th-grade cohorts.⁷ The 20,422 students are evenly distributed across cohorts such that 25% are from each cohort.

Data on students are included in the NSC dataset if they enrolled in a college that is included among the 93% of colleges whose admissions data is collated in the Clearinghouse at some point during the first four semesters post-graduation from high school. Fifty percent of the 20,422 students in the CPS sample are located in the NSC dataset, indicating that at least half of CPS students in the full sample enrolled in two- or four-year college for at least one semester during the first four semesters after their graduation from high school.

SAMPLE

The full sample of CPS students ($n=20,422$) contains every 11th-grade student in each of the 13 CPS high schools that offered the IB program to the 11th-grade cohort in academic years 2002-2003 ($n=5,042$),

2003-2004 (n=5,097), 2004-2005 (n=5,181), and 2006-2007 (n=5,102). The study uses data from these years because college enrollment data is available on each of these cohorts. Four annual cohorts of students are included to improve the statistical power of the analyses and enhance the external validity of the findings.

MEASURES

Table 1 presents descriptive statistics on the outcomes, question predictors, and covariates for the full sample of students included in the CPS dataset (n=20,422). The first outcome variable is a measure of student achievement, formed by averaging 11th-grade students' mathematics, English, reading, and science scores on the ACT college entrance examination.⁸ In total, 2,667 students—1% of IB students and 14% of non-IB students—are missing on at least one of the four constituent achievement tests and thus are also missing on the academic achievement outcome itself. After compositing, for ease of interpretation, I standardized the ACT composite to a mean of zero and a standard deviation (SD) of unity in the analytic sample of students (n=14,368) who were not missing seventh-grade mathematics or reading scores nor one of the ACT subject-level tests. Students take the constituent achievement tests during the late spring of their junior year, and so estimated “treatment effects” represent the impact of only one year of official IB enrollment on the academic-achievement outcome.

The second outcome is a dichotomous indicator that is coded 1 if students graduated on time from high school with a regular CPS degree or an alternate Illinois degree (0 otherwise).⁹ This information is included in the data for every student in the full sample, of which 99% of IB and 78% of non-IB students graduated from high school.¹⁰

The third outcome is a dichotomous indicator that is coded 1 for students who enrolled in either two- or four-year college for at least one semester during one of the first four semesters post-graduation from high school (0 otherwise). Among IB students, 83% enrolled in college compared to 48% of non-IB students. The NSC only matches CPS high school graduates to its database. Therefore, students' college enrollment status is conditional on whether they graduated from high school. If IB enrollment affects the probability that students will graduate from high school, then relative to non-IB students, a greater proportion of IB students will be included in the sub-sample of students used to estimate the impact of IB enrollment on students' probability of college enrollment. The overrepresentation of IB students would then contribute to biasing the estimation of the impact of IB enrollment on students' probability

Table 1: Descriptive Statistics for Selected Variables on the Full Sample of 11th-Grade Students Who Enrolled in the 13 CPS Schools That Offered IB to the 2002-2003, 2003-2004, 2004-2005, and 2006-2007 11th Grade Cohorts (n=20,422).

Variable descriptions	IB students		Non-IB students	
	Mean	S.D.	Mean	S.D
<i>Outcomes:</i>				
Took the ACT test	0.992		0.860	
Linear composite of the four ACT subject matter tests	1.488	1.134	-0.126	0.878
Graduated from high school	0.989		0.767	
Enrolled in college	0.830		0.481	
<i>Question predictor:</i>				
Enrolled in IB	Full sample mean=0.069			
<i>Covariates:</i>				
Seventh-grade ITBS mathematics percentile	79.468	16.180	50.269	24.950
Whether students are missing their seventh-grade math percentiles	0.199		0.262	
Seventh-grade ITBS reading percentile	73.855	17.707	47.128	22.974
Whether students are missing their seventh-grade reading percentiles	0.200		0.261	
Male		0.350		0.477
Eligible to receive free or reduced-price lunch	0.647		0.778	
Whether students are Asian American	0.165		0.058	
Whether students are African American	0.280		0.370	
Whether students are Latino	0.297		0.421	
Whether students are Native American/Alaskan	0.004		0.002	
Whether students are White	0.255		0.149	

of college enrollment. In the “Threats to Validity” section of this paper, I discuss this potential threat to a causal interpretation of the estimates of the impact of IB enrollment on students’ probability of college enrollment.

The question predictor (*IBTREAT*) is defined as whether students enrolled in the core IB Theory of Knowledge course for at least a semester (1=yes, 0=no). Unlike alternative ways of defining the IB treatment, such as whether students sit for IB examinations or earn IB subject-specific

diplomas, this definition captures IB enrollment rather than IB persistence.¹¹ According to the CPS IB Diploma Program administrator (personal correspondence, July 6, 2010), students typically only enroll in TOK if they are attempting to fulfill the IB requirements and only rarely as a way to take honors courses in a piecemeal fashion. The students that enroll in TOK, on average, enroll in 11 IB courses ($SD=4$), confirming this characterization of TOK enrollment. Per this definition, seven percent (1,403 students) of the full sample of 20,422 students was enrolled in the IB treatment. The percentages of IB students per cohort were similar across cohorts at 6% in 2003 and 2004, 7% in 2005, and 9% in 2007.

The analysis includes students' seventh-grade mathematics and reading test percentiles, which are based on their national percentile ranks on the ITBS, as measures of their previous academic achievement. Professors from the University of Iowa developed and refine the tests—intended to diagnostically determine students' academic progress and drive curricular decisions—and run the Iowa Basic Skills Testing Program. Each subject matter test takes students 30 minutes or less to complete and is marked by hand or through local or centralized scoring options. Reading tests administered to seventh graders include vocabulary, reading comprehension, spelling, capitalization, punctuation, and usage items. Seventh-grade mathematics test items address concepts and estimation, problem solving, and data interpretation and computation (Riverside Publishing, 2010). Sample students took the ITBS tests in the spring of seventh grade and CPS administrators used these test percentiles to determine eighth graders' eligibility for ninth-grade IB enrollment. In the full sample, 26% (5,261) of students are missing these scores, compared to 20% (279) of IB students. In the "Threats to Validity" section, I also discuss this potential threat to a causal interpretation of the estimates of the impact of IB enrollment. I eliminate students missing their seventh-grade test percentiles, leaving an analytic subsample of 15,148 students. Table 1 shows that among IB students, the mean ITBS mathematics-achievement percentile is 79.47 ($SD=16.18$) and the mean reading-achievement percentile is 73.86 ($SD=17.71$). In comparison, the non-IB mean mathematics percentile of 50.27 ($SD=24.95$) and mean reading percentile of 47.13 ($SD=22.97$) are each almost 30 percentiles lower. Across cohorts, ITBS mathematics and reading achievement percentile averages among IB and non-IB students, respectively, remain fairly constant.

Other covariates include measures of students' demographic characteristics, including their gender, family income, and race. A parent or guardian designates his or her child's gender on the CPS enrollment

form (Illinois State Board of Education, 2013). Male equals 1 for the boys, who make up 48% of non-IB students and 35% of IB students (*MALE*). The gender breakdown was highly consistent across cohorts for both IB and non-IB students.

Family income is represented with a dichotomous indicator that is coded 1 for students who were eligible to receive free or reduced-price lunch. Illinois schools, including CPS, currently participate in the National School Lunch Program (NSLP), and did during the study time-frame. According to NSLP regulations, there are two main ways through which free or reduced-price lunch eligibility is determined: (1) Families have the opportunity to self-report their eligibility using the Household Eligibility Application, and (2) Through the Direct Certification method, the Illinois State Board of Education reports to school districts all children who are part of families eligible for food stamps and/or the Temporary Assistance for Needy Families program, and these children are eligible for the lunch program (U.S. Department of Agriculture, 2012). Across cohorts, 65% of IB students were eligible, compared to 78% of non-IB students, indicating that on average, both non-IB and IB students are from low-income families, though the probability that non-IB students are from poorer families is greater. The free or reduced-price lunch breakdown is also very consistent by IB status across cohorts, with a minimum of 61 percent of eligible IB students in 2005 and a maximum of 68 percent of eligible IB students in 2004.

Student race is represented with a vector of dichotomous indicators that describe whether a student is Asian American, African American, Latino, Native American or Alaskan American, or White, dropping the last designation to create a reference category. These five race and ethnicity reporting categories were those specified by the U.S. Department of Education prior to 2010. States were required to use these categories to report aggregate data to the federal government (the categories changed in 2010). In most cases, parents or guardians report their child's race/ethnicity using a form created by the local school district that is based on a state template form. If parents decline to indicate race/ethnicity, school and/or district staff is required to do so (Illinois State Board of Education, 2009). Overall, most IB and non-IB students are minority and their distributions across Latino, African-American, White, and Asian-American racial groups are fairly comparable. Race proportions among IB and non-IB students are also consistent across cohorts.

The generic variable name *STUDENTDEM* represents the vector of covariates that record students' ITBS test percentiles, gender, family income, and race.

ANALYTIC METHODS

The objective of this study is to estimate the causal impact of enrollment in the IB Diploma Program on CPS 11th-grade students' academic achievement, probability of high school graduation, and probability of college enrollment. The ideal way to estimate the causal impact of the IB Diploma Program enrollment on students' outcomes would have been to exploit a situation in which for some exogenous reason students were assigned randomly to participate in an IB Diploma Program. In theory, the CPS test-score rule used to determine IB eligibility provides this opportunity. In practice, however, low student take-up in IB among students eligible to enroll by virtue of their seventh-grade scores makes it infeasible to exploit this potential source of exogenous variation. In the absence of other sources of random variation, this study uses a *propensity-score approach* in combination with sensitivity analyses to demonstrate a causal relationship between IB Diploma Program enrollment and student outcomes.

In this study, the same analytic strategy addresses each research question, described for outcome Y , for student i in high school j in cohort c . The first step is to assume selection into the IB Diploma Program based on students' observed characteristics by using a propensity-score approach (described in detail below) to estimate the impact of IB enrollment on three measures of students' academic success. Then, as discussed in the "Threats to Validity" section of this paper, following Rosenbaum (2002), the second step is to test the sensitivity of the estimates to different levels of selection bias.

Compared to using Ordinary Least Squares (OLS) methods to address the research questions, the propensity-score approach used in this study is a more effective, robust, and partially non-parametric way of controlling for observed bias attributable to measured covariates (Murnane & Willett, 2011, p. 310). The propensity-score approach requires first modeling and then examining explicitly the process of students' selection into the IB program. The extent to which the selection model successfully captures the selection process improves the ability to remove selection bias from subsequent estimates of the causal impact of IB enrollment on academic outcomes (Murnane & Willett, 2011, p. 288). In addition, *propensity scores* constitute the basis of Rosenbaum's sensitivity analysis, to which the results are subjected and described later in the "Threats to Validity" section.

The propensity-score weighting strategy described below—as compared to stratifying or matching strategies—is the preferred use of the propensity scores because it permits the imposition of identical (balanced)

covariate distributions for IB and non-IB groups and to include in the comparison sample all non-IB students who are most comparable to IB students based on observed characteristics (Murnane & Willett, 2011).

In fitting the *propensity-score selection model*, I use logistic regression analysis to model students' probability of enrolling in the IB program as a function of a set of measured characteristics that theoretically should predict enrollment (Murnane & Willett, 2011, p. 311). Students' seventh-grade mathematics and reading test percentiles are included as the principal predictors in the propensity-score selection model because these test results should be strong predictors of whether students enroll in the IB Diploma Program, given the CPS IB-admission policy.¹² The model also includes an indicator of gender because CPS boys' tendency to fail classes, earn lower grade-point averages, and drop out of school at higher rates than girls (Allensworth, 2005) suggests that they have a lower theoretical probability of CPS IB enrollment eligibility than girls. The model includes measures of race and family income as predictors because historically IB has attracted White, Asian-American, and economically advantaged students, while African-American, Latino, and economically disadvantaged students have not had the academic preparation necessary to enter the rigorous Diploma Program (Burris, Welner, Wiley, & Murphy, 2007).

I therefore fit the following logistic regression model (1) to represent the probability that student i in school j in cohort c chooses to enroll in the IB program:

$$(1) \Pr(IBTREAT_{ijc} = 1) = \frac{1}{1 - e^{-\beta_0 + \beta_1 STUDENTDEM_{ijc} + \delta_j + \lambda_c}}$$

Included school (δ_j) and cohort (λ_c) fixed effects allow for the probability that student i will enroll in IB to differ by school and cohort. The fixed effects address the concern that schools' and cohorts' unmeasured characteristics might contribute to determining whether students will enroll in IB.

Based on the fitted propensity-score selection model, the propensity scores (*PHAT*) are the predicted probabilities that students will enroll in the IB program. After obtaining *PHAT* for each of the students in the analytic sample, I follow Imbens and Wooldridge (2009) in eliminating from the analytic sample all non-IB students with estimated propensity scores that fall outside of the range of estimated propensity scores for the IB students. This process aligns the IB and non-IB groups explicitly on the selection model covariates, ensuring that the non-IB students in the comparison sample are as similar as possible in terms of measured

characteristics to IB students. The resulting comparison sample includes 11,592 student records.

Following Hirano, Imbens, and Ridder (2003), *PHAT* provides an inverse-probability weighting, *IPW*, for each student in the comparison sample. The weights are created as follows: For IB students: $IPW=1/PHAT$; for non-IB students: $IPW=1/(1-PHAT)$. The purpose of this weighting strategy is to allocate the least weight to the IB students with the highest propensity for selection into IB and the greatest weight to the IB students with the lowest propensity for selection into IB (and vice versa for non-IB students) in order to impose identical (observed) covariate distributions for students who enrolled versus those who did not enroll in the IB program. The findings described in the Results section and shown in Table 3 demonstrate that weighted with their *IPW*, IB and non-IB students in the comparison sample ($n=11,592$) are balanced on all covariates.

The final step is to fit a weighted least squares (WLS) estimation model (2), which represents the relationship between my academic outcomes and membership in the IB treatment, weighting each student i in school j in cohort c by his or her *IPW* as follows:¹³

$$(2) Y_{ijc} = \beta_0 + \beta_1 IBTREAT_{ijc} + \beta_2 STUDENTDEM_{ijc} + \delta_j + \gamma_c + \epsilon_{ijc}$$

Where Y is the outcome in question. The model controls for the vector of student characteristics, *STUDENTDEM*, and includes school (δ_j) and cohort (γ_c) fixed effects to increase the precision of the estimates (Imbens & Wooldridge, 2009). Residuals are represented with the term E and standard errors are estimated to account for covariation in students' outcomes within schools, the unit at which the IB treatment is administered. Assuming selection into IB based only on students' observed characteristics, β_1 can be interpreted as the impact of IB enrollment on outcome Y .

To determine if the impact of IB enrollment differs by gender, I fit model (3) below, including the main effect of *MALE* in the *STUDENTDEM* term, the two-way interaction term *IB*MALE* and again weighting each student i in school j in cohort c by *IPW*:

$$(3)$$

$$Y_{ijc} = \beta_0 + \beta_1 IBTREAT_{ijc} + \beta_2 IB*MALE_{ijc} + \beta_3 STUDENTDEM_{ijc} + \delta_j + \gamma_c + \epsilon_{ijc}$$

The β_1 parameter estimate is interpreted as the impact of IB enrollment on girls and the β_2 parameter estimate as the extent to which the impact of enrollment in IB differs according to gender.

MISSING DATA SUMMARY

The full sample includes 20,422 CPS students. The 3,469 students who were missing on an IB treatment status are defined as being non-IB rather than dropped. The rationale is that according to model (1), the distribution of propensity scores for students missing on IB status (mean=0.053, SD=0.113) is to the left of the distribution of propensity scores for students who explicitly did not choose the IB option (mean=0.065, SD=0.113). Therefore, students missing on IB status, if anything, look more like non-IB students than IB students, for whom the mean propensity score is 0.294 (SD=0.210).

Since seventh-grade test scores are critical to the propensity model, I eliminate students missing their seventh-grade test percentiles, leaving an analytic subsample of 15,148 students. Then, following Imbens and Wooldridge (2009), I eliminate from the analytic sample all non-IB students with estimated propensity scores that fall outside of the range of estimated propensity scores for the IB students ($n=3,556$), which leaves a resulting comparison sample of 11,592 student records. Of the 11,592 students in the comparison sample, 22% are from the 2003 cohort, 25% from 2004, 25% from 2005, and 28% from 2007. As mentioned above, of the full sample of 20,422 students, 6% of the 2003 and 2004 cohorts, 7% of the 2005 cohort, and 9% of the 2007 cohort were IB students. Given that a key aspect of the analytic strategy is to ensure that IB and non-IB students in the comparison sample are as similar as possible, it makes sense that the comparison sample is composed of a slightly higher percentage of 2007 cohort students compared to those from the 2003 cohort.

RESULTS

THE IB ENROLLMENT PROCESS

Table 2 shows the results of fitting *Model (1)*, which describes the fitted relationship between students' probability of enrolling in the IB Diploma Program and predictors that include their demographic characteristics and previous academic achievement. The table presents parameter estimates and standard errors for each of the demographic and ITBS test percentile covariates that participate in the IB selection process.

In addition, to demonstrate the substantive nature of the IB selection process visually, Figure 1 plots the fitted probability that prototypical students will enroll in the IB program as a function of their seventh-grade ITBS reading percentiles. Fitted lines represent students who attended one CPS high school that offered the IB program to the 11th-grade

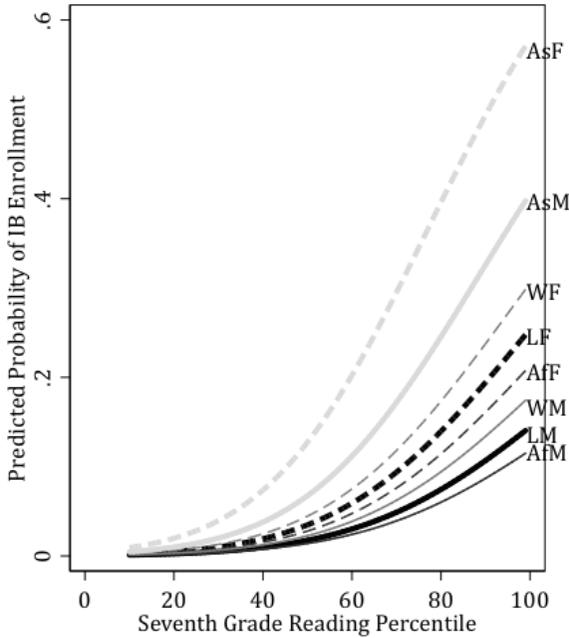
Table 2. Fitted Logistic Regression Model (Standard Errors) and Approximate P-Values Describing the Prediction of Students' Probability of Enrolling in the Ib Program, Based on Their Demographic Characteristics and Previous Academic Achievement (n=15,148)

Predictors	Parameter Estimate (Standard error)
Male	-0.701*** (0.103)
African American	-0.485** (0.240)
Latino	-0.256 (0.220)
Asian American	1.140*** (0.175)
Alaskan & Native American	-1.251 (1.039)
Free/reduced-price lunch	-0.210 (0.158)
Seventh-grade mathematics score	0.034* -0.02
Seventh-grade mathematics score ^ 2	0.0001 (0.0002)
Seventh-grade reading score	0.084*** (0.024)
Seventh-grade reading score ^ 2	-0.0003 (0.0002)
Constant	-9.189*** (1.677)

*** p<0.01, ** p<0.05, * p<0.1

2006-2007 cohort, who scored at the 50th percentile on the ITBS mathematics test, and who were eligible to receive free or reduced-price lunch (n=303). The eight fitted lines compare the fitted probabilities that Latino, African-American, White, and Asian-American girls and boys enrolled in the IB program, as a function of their seventh-grade reading test percentiles. Solid lines represent the fitted selection probabilities for boys and dashed lines for girls. Shades of gray and white represent each racial group.

Figure 1. Predicted probability of IB enrollment by prior seventh-grade reading percentiles, for students of different genders and races (n=303)



AfF=African-American Female
AfM=African-American Male
AsF=Asian-American Female
AsM=Asian-American Male
LF=Latino Female
LM=Latino Male
WF=White Female
WM=White Male

Among Title 1 U.S. high schools that offer the IB program, approximately two-thirds require that students demonstrate strong previous academic ability to be eligible to enroll. The most commonly used selection criterion is grade point average, followed by standardized test performance (Siskin & Weinstein, 2008). Like enrollment policies in the majority of IB-offering Title 1 schools, the CPS IB enrollment policy requires that students demonstrate their academic ability to be eligible to enroll. Specifically, students must score over the 50th percentile on both the mathematics and reading seventh-grade tests to be eligible. Given this enrollment requirement, I had anticipated and the CPS IB Diploma Program administrator concurred that it would be more probable that

students who have demonstrated stronger prior academic achievement would enroll in the IB Diploma Program (personal correspondence, July 6, 2010).

Among the highlights of the results shown in Table 2, it is evident that, holding constant gender, race, lunch-eligibility status, school, and cohort, the fitted probability that students will enroll in the IB program is greater for students with higher seventh-grade mathematics ($p=0.087$) and reading ($p=0.001$) scores. Correspondingly, Figure 1 shows that for all students who score below the 50th percentile on the seventh-grade ITBS reading score, there is close to a zero fitted probability of IB enrollment.

Gender is also an important predictor of IB enrollment. As shown in Table 2, holding constant seventh-grade test scores, race, lunch-eligibility status, school, and cohort, the fitted odds that a boy will enroll in the IB program is half that of a girl ($e^{-.701}=0.50$; $p<0.001$). In Figure 1, within racial groups represented by the same color line, each of the dashed fitted lines (representing prototypical girls) indicates that there is a greater fitted probability of enrollment in the IB program than do the solid fitted lines representing boys. In fact, as the dashed lines show, the fitted probability that all White, Latino, and African-American girls will enroll in the IB program is greater than the probability that all White, Latino, and African-American boys will enroll.

The race-based differential in the fitted probability of enrollment in the IB program aligns with the Burris et al. (2007) argument that White and Asian-American students, on average, are better prepared for rigorous courses like those offered through IB than are African-American and Latino students. In the CPS context, the fitted lines in Figure 1 show that, holding constant seventh-grade test scores, lunch eligibility status, school, and cohort, the fitted probability that Asian-American students—boys and girls—enroll in the IB program is greater than the corresponding probability that students of any other racial group enroll ($p<0.001$). In addition, the fitted probability of White and Asian-American students' IB enrollment is greater than the corresponding probability for African-American and Latino students' enrollment.

Table 3 shows the differences in measured covariates between IB and non-IB students, unweighted and then weighted, for the comparison sample of students with overlapping propensity scores ($n=11,592$).

Though the comparison sample is restricted to include only non-IB students with propensity scores that overlap with those of IB students, statistically significant differences on covariates between the two groups remain when the analyses do not incorporate the *IPW* weights. As the third column shows of Table 3 shows, unweighted, non-IB students are not comparable to the IB students on gender or race.

Table 3. Unweighted and Weighted Differences Between non-IB and IB Student Demographics and Test Scores (n=11,592)

Characteristics	Unweighted Means and Differences			Weighted Means and Differences		
	IB=0	IB=1	Difference	IB=0	IB=1	Difference
	Mean (SD)	Mean (SD)	(SE)	Mean (SD)	Mean (SD)	(SE)
	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.426 (0.495)	0.343 (0.475)	-0.083** (0.029)	0.419 (0.493)	0.406 (0.491)	-0.013 (0.042)
African American	0.407 (0.491)	0.304 (0.460)	-0.103* (0.050)	0.397 (0.489)	0.353 (0.478)	-0.044 (0.063)
Latino	0.424 (0.494)	0.337 (0.473)	-0.087 (0.056)	0.415 (0.493)	0.382 (0.486)	-0.033 (0.052)
Asian American	0.046 (0.210)	0.151 (0.359)	0.105** (0.043)	0.058 (0.233)	0.094 (0.292)	0.036 (0.027)
Alaskan/Native	0.002 (0.041)	0.0009 (0.030)	-0.0008 (0.001)	0.002 (0.040)	0.0006 (0.024)	-0.001 (0.0008)
Free/reduced-price lunch	0.783 (0.412)	0.681 (0.466)	-0.102 (0.058)	0.774 (0.418)	0.824 (0.380)	0.050 (0.040)
ITBS reading	54.681 (19.113)	73.855 (17.707)	19.174*** (2.081)	56.664 (19.85)	55.024 (23.26)	-1.164 (4.183)
ITBS reading ^ 2	3355.3 (2454.4)	5767.8 (2454.4)	2412.6*** (328.3)	3604.7 (2346.9)	3568.2 (2528.1)	-36.5 (399.2)
ITBS mathematics	59.641 (20.223)	79.468 (16.187)	19.827*** (1.996)	61.675 (20.705)	61.286 (21.514)	-0.389 (3.260)
ITBS mathematics ^ 2	3965.965 (2369.2)	6576.9 (2344.1)	2610.9*** (325.8)	4232.5 (2484.5)	4218.4 (2606.3)	-14.098 (383.9)

Notes. The sample (n=11,592) is the same in unweighted and weighted columns and includes all students in the comparison sample who are not missing on their seventh-grade test score percentiles and who have overlapping propensity scores.

*** p<0.01, ** p<0.05, * p<0.1

However, the sixth column of Table 3 shows that after eliminating non-comparable, non-IB students from the comparison sample, *and* weighting students by their *IPW*, differences in the distributions of observable characteristics between IB and non-IB students disappear. The weights impose identical covariate distributions on both the IB and non-IB students such that when estimating differences in outcomes between *IPW*-weighted IB and non-IB students in subsequent WLS regression analyses, the groups are indistinguishable based on measured characteristics.

Weighted, the groups are also indistinguishable on all covariates by cohort, with two minor exceptions: Compared to non-IB students, there is a lesser sample probability that IB students in the (1) 2003 cohort will be Latino and (2) 2005 cohort will be Alaskan or Native American.

RQ #1: THE IMPACT OF IB ENROLLMENT ON STUDENTS' ACADEMIC ACHIEVEMENT

Table 4 presents the *IPW*-weighted least-squares fitted regression models. The columns show four separate fitted models that use the propensity-score weighting approach to predict the impact of IB enrollment on students' (1) probability of taking the ACT test; (2) academic achievement as measured by the variable *ACT*; (3) probability of high school graduation; and (4) probability of college enrollment. Each column shows the outcome's mean and SD in the comparison sample and then the parameter estimate and associated standard error of the impact of IB enrollment, by outcome.

Since 2002, Illinois has strongly encouraged and paid for all students to take the ACT college admissions examination (Illinois Department of Education, 2004). The ACT organization reports that annually, approximately 98% of Illinois students take the four, subject-specific ACT test sections. Among students in the comparison sample, 96% took the four test sections, as indicated by having their *ACT* measure information. The first column of Table 4 shows that enrollment in the IB Diploma Program increases—by three percentage points ($p=0.009$)—the fitted probability of taking the ACT examination. The second column shows that, on average, accounting for measured characteristics that affect students' selection into the IB program, IB enrollment increases students' academic achievement as measured by the ACT examination by 0.54 SDs ($p<0.001$) or about one half of a SD. The Appendix table shows that the respective IB enrollment impacts on the fitted probability of taking the ACT examination and on ACT scores is comparable across cohorts.

Having ACT scores is a necessary condition for inclusion in the sub-sample of students that is used to estimate the impact of IB enrollment on students' ACT scores. That IB enrollment increases students' probability of taking the ACT examination may affect the composition of unobserved characteristics among the sub-sample of IB and non-IB students who took the ACT examination, and thus may contribute to biasing the estimation of the impact of IB enrollment on students' ACT scores. I discuss this potential threat to the causal interpretation of the estimate of the impact of IB enrollment on students' ACT scores in the following "Threats to Validity" section.

Table 4. Parameter Estimates (Standard Errors) and Approximate P-Values for Four Separate Wls-Fitted Models That Describe the Impact of Ib Enrollment on Four Student Outcomes

	Outcome Variables			
	(1)	(2)	(3)	(4)
	Takes the ACT Examination	Academic Achievement (ACT Scores)	Probability of High school Graduation	Probability of College Enrollment
Outcome mean	0.960	0	0.829	0.585
(SD)	(0.195)	(1)	(0.376)	0.493
Predictor				
Enrolled in IB	0.034*** (0.011)	0.536*** (0.052)	0.174*** (0.018)	0.217*** (0.024)
Observations	11,592	11,148	11,592	11,592
R ² Statistic	0.032	0.692	0.068	0.127

Notes. Samples include all students not missing on their seventh-grade test scores and non-IB students with propensity scores that overlap with those of IB students.

*** p<0.01, ** p<0.05, * p<0.1

RQ #2: THE IMPACT OF IB ENROLLMENT ON STUDENTS' PROBABILITY OF HIGH SCHOOL GRADUATION

Among the students in the comparison sample, 83% graduated from high school. Column three of Table 4 shows that IB enrollment increases the probability that students will graduate from high school by 17.4 percentage points ($p < 0.001$) or by 20% (0.17/0.83). The Appendix table shows that the IB enrollment impact on the probability of high school graduation is comparable across cohorts.

RQ #3: THE IMPACT OF IB ENROLLMENT ON STUDENTS' PROBABILITY OF COLLEGE ENROLLMENT

Of the students in the comparison sample, 59% enroll in college during at least one of the first four semesters post-graduation from high school. The fourth column of Table 4 shows that IB enrollment increases students' probability of enrolling in college by 21.7 percentage points ($p < 0.001$) or 38% (0.22/0.59). The Appendix table also shows that the IB enrollment impact on the probability of college enrollment is comparable across cohorts.

Table 5. Parameter Estimates (Standard Errors) and Approximate P-Values for Models (1)-(4), for Wls-Weighted Analyses of the Impact of Ib Enrollment on Four Student Outcomes

	Outcome Variables			
	(1)	(2)	(3)	(4)
	Takes the ACT Examination	Academic Achievement (ACT Scores)	Probability of High school Graduation	Probability of College Enrollment
Predictors				
IB impact for girls	0.021** (0.009)	0.464*** (0.053)	0.138*** (0.014)	0.194*** (0.028)
Difference in IB impact by gender	0.035** (0.012)	0.190*** (0.034)	0.095*** (0.018)	0.059** (0.024)
Observations	11592	11148	11592	11592
R ² Statistic	0.034	0.693	0.071	0.127

Notes. Other terms, including the main effects of gender, are present in the fitted models. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

RQ #4: GENDER DIFFERENCES IN THE IMPACT OF IB ENROLLMENT

Table 5 shows the results of the analyses of whether there are gender-based differences in the IB enrollment impact. The columns present four separate fitted models that use propensity-score weighting to estimate the impact of IB enrollment. Each row, separately for the models (1)-(4), presents parameter estimates and standard errors for the impact of IB enrollment for girls and for the extent to which the impact of enrollment in IB differs according to gender.

Table 5 shows that the impact of IB enrollment on students' outcomes is greater for boys than for girls for all outcomes. The first column shows that among IB students, the probability that boys will take the ACT is 3.5 percentage points greater than the probability for girls ($p = 0.014$). Next, among IB students, on average, boys demonstrate academic achievement of a magnitude that is 0.19 SDs higher than girls ($p < 0.001$). The third column shows that compared to female IB students, the probability that male IB students will graduate from high school is 9.5 percentage points greater ($p < 0.001$). Finally, the probability that male IB students will subsequently enroll in college is 5.9 percentage points greater ($p = 0.034$).

THREATS TO VALIDITY

CHALLENGE #1: SELECTION BIAS

Unlike an experimental evaluation of the IB treatment, students were not randomly assigned to a treatment (IB) or a control (regular high school) condition. On the contrary, students who *chose* to participate in an IB program may differ from students who do not, in important measured and unmeasured ways that may also be related to students' subsequent academic outcomes. Ignoring the effects of these unmeasured characteristics in the analyses will lead to *selection bias* in estimates of the effect of enrollment in the IB program on students' academic outcomes. For example, assuming that students who choose to enroll in IB are more motivated and/or harder-working—or that their parents are more knowledgeable about how to navigate different high school options—and that these unmeasured characteristics lead them to perform better on measures of academic success, then estimates of these effects will overestimate the IB enrollment impact.¹⁴

The ideal way to estimate the impact of IB enrollment on students' academic outcomes would be to assign students randomly to participate in IB, or not, and to compare the outcomes of students enrolled in the IB program to those of students who are not. This study, in the absence of random assignment data, uses observational data to estimate the IB enrollment impact. Since the propensity-score approach estimates shared above are likely overestimates of the true impact of enrollment in the IB program, I assess the sensitivity of my estimates of effect to the possibility of differing degrees of selection bias.

The sensitivity analyses, following Rosenbaum (2002), ask how much selection bias would need to exist in order to negate the impacts of IB enrollment detected with the propensity-score approach. The strategy that Rosenbaum has suggested for assessing the sensitivity of the findings relies on Wilcoxon's rank test and is appropriate for continuous outcomes like the measure of academic achievement. For dichotomous outcomes, I conduct similar sensitivity analyses by adopting the same strategy, but base these analyses instead on McNemar's (1947) test, which compares matched IB and non-IB student pairs who have dissimilar values of a dichotomous outcome.¹⁵

Under Rosenbaum, the symbol *gamma* (Γ) represents different levels of selection bias. In a matched, randomized experiment, two individuals within a matched pair (with identical covariates) have the same pre-treatment, or *ex-ante*, probability of receiving treatment. In this case, (Γ)=1. In an observational study, two individuals with identical covariates might

have different *ex-ante* probabilities of treatment because an unobserved factor (or combination of factors)—like motivation—that potentially affects both the probability of treatment and subsequent outcomes, might make the probability that one person receives the treatment greater. If, based on the unobserved factor(s), the *ex-ante* probability that one person receives the treatment were twice the probability that the other receives the treatment, Γ would equal 2. Following Rosenbaum, I estimate how high Γ must be to negate completely the estimate of a statistically significant difference in findings between IB and non-IB pairs.

The first step is to assume several different reasonable levels of Γ . The next, based on the appropriate Wilcoxon rank test or McNemar test, is to estimate respective *p*-values associated with upper and lower bounds around the distribution of the estimate of the IB impact on a given outcome at a given level of Γ . By examining whether the upper bound *p*-value indicates to reject the null hypothesis of no impact of IB enrollment on a given outcome, I can assess the extent to which selection bias affects the findings. When the *p*-value associated with the upper bound on the distribution of the estimated IB impact is greater than 0.05 at *lower* levels of Γ , the conclusion is that the estimate of the IB program effect is sensitive to selection bias, indicating that much of the estimated impact could be attributable to unobserved differences in students' characteristics between the IB and non-IB groups. Conversely, when the upper bound *p*-value is greater than 0.05 only at *higher* levels of Γ , the conclusion is that the estimated IB enrollment impact is less sensitive to selection bias, and that much of the estimated impact can be attributed to IB enrollment. Keele (2009) states that in the social sciences, most findings are typically robust to values of Γ between 1 and 2.

Rosenbaum (2002) also demonstrates that another way to understand Γ is as another variable in the propensity-score selection model, but one that predicts enrollment in the IB Diploma Program based on unobserved characteristics like motivation, rather than an observed characteristic like gender. For example, an interpretation of the results shown in Table 2 is that the odds that a boy will enroll in the IB program is half that of a girl ($e^{-.701}=0.50$). A value of Γ greater than 2 implies that a covariate representing unmeasured characteristics would need to multiply the odds of IB enrollment by more than four times ($2/0.50$) that of gender to completely invalidate any positive impact of IB enrollment on a given measure of students' academic achievement.

Table 6 presents the results of the sensitivity analyses for the impact of IB enrollment on students' academic achievement, as measured by the continuous outcome ACT. In the table, the phrases "Lower" and "Upper" represent the lower and upper Hodges-Lehmann point estimates

Table 6. Rosenbaum Sensitivity Analyses Displaying the Extent to Which Propensity Score Estimates of the Impact of Ib Enrollment on Students’ Academic Achievement Are Sensitive to Different Levels of Selection Bias, Represented by Γ (1,116 Matched Pairs)

Γ	Lower	Upper
1	0.499	0.667
2	0.075	1.100
2.2	0.017	1.157
2.3	-0.009	1.185

associated with the upper-bound p -value implied by a given level of Γ .0

In the first row of Table 6, assume as if analyzing data from a randomized trial, that $\Gamma=1$. If this were the case, the point estimate of the impact of IB enrollment on students’ academic achievement would fall between a lower bound point estimate of 0.499 ACT examination SDs and an upper bound point estimate of 0.667 SDs. Assuming that $\Gamma=2$, as shown in the second row, the same point estimate would fall within a range of 0.075 to 1.100 SDs. Note that neither of these intervals cover zero. The third row shows that the highest level that Γ can attain if the lower and upper bounds do not include zero is 2.2. Thus, conditional on all the observed covariates included in the models, *ex-ante*, the probability that an IB student will enroll in IB could be 2.2 times—or 220%—the probability that a non-IB student will enroll, and the effect of IB enrollment on students’ academic achievement will still be statistically significant. Returning to Rosenbaum’s explanation, this value of Γ implies that selection bias would need to multiply the odds of IB enrollment by more than 4.6 times (2.3/0.50) that of gender to completely invalidate any positive impact of IB enrollment on students’ ACT scores.

Panel A of Table 7 shows that the estimate of the impact of IB enrollment on students’ probability of graduation from high school is highly insensitive to selection bias. Even assuming that $\Gamma=10$, the associated p -value of 0.005 indicates rejection of the null hypothesis of no impact of IB enrollment on students’ probability of high school graduation. The fifth row shows that *ex-ante*, the probability that an IB student will enroll in IB would have to be 13 times the probability that a non-IB student will enroll, in order to negate the significance of my finding that IB enrollment increases students’ probability of high school graduation. To contextualize a Γ of 13, Rosenbaum describes a Γ of 6 as “enormous” in his test of the sensitivity to selection bias of his estimated impact of smoking on suffering from lung cancer (2002, p. 79).

Panel B of Table 7 shows that the p -value associated with a Γ of 2.6 is greater than 0.05. *Ex-ante*, selection bias would need to make the

Table 7. Rosenbaum Sensitivity Analyses Displaying the Extent to Which Propensity Score Estimates of the Impact of Ib Enrollment on Students’ Probability of High School Graduation (Panel A) and Probability Of College Enrollment (Panel B) Are Sensitive to Different Levels Of Selection Bias, Represented by Γ (1,123 Matched Pairs)

Panel A: Probability of High school Graduation	
Γ	p-value
1	1.664e-49
5	1.857e-08
10	0.005
12.5	0.042
13	0.057
Panel B: Probability of College Enrollment	
Γ	p-value
1	5.203e-30
2	0.00003
2.5	0.039
2.6	0.083

probability that IB students will enroll in IB 2.6 times—or 260%—the probability that non-IB students will enroll, in order to negate completely the finding that IB enrollment increases the probability of college enrollment. Selection bias would need to multiply the odds of IB enrollment by more than 5.2 times (2.6/0.50) that of the effect of gender to completely invalidate any positive impact of IB enrollment on probability of college enrollment.¹⁶

CHALLENGE #2: IB ENROLLMENT INCREASES THE PROBABILITY THAT STUDENTS TAKE THE ACT EXAMINATION

In addition to the challenge posed by selection bias, another internal-validity threat is that enrollment in the IB Diploma Program increases the fitted probability of taking ACT examination by three percentage points ($p=0.009$). To understand why this finding may contribute to underestimating the true impact of IB enrollment on students’ ACT scores, I make one assumption: On average, students who took the ACT test scored higher than students who did not take the test would have—if they had taken the test. This assumption is plausible because, given Illinois’s universal ACT policy, the most likely reason for not taking the ACT examination is student absenteeism, and higher rates of student absenteeism tend to be associated with lower achievement (Marburger, 2006).

Given this assumption, if the same proportion of IB and non-IB students took the ACT examination, then the inclusion of the ACT scores of the non-IB students who would not otherwise have taken the ACT test would diminish the average ACT test score among non-IB students. This lower average non-IB ACT score would result in a greater estimated average difference in ACT scores between IB and non-IB students.

CHALLENGE #3: IB ENROLLMENT INCREASES THE PROBABILITY THAT STUDENTS GRADUATE FROM HIGH SCHOOL

The third internal validity challenge stems from estimating the impact of IB enrollment on students' probability of college enrollment, when college enrollment is conditional on high school graduation and IB enrollment increases the fitted probability of graduating from high school. Based on the plausible assumption that the unmeasured characteristics that motivate students to graduate from high school also motivate them to enroll in college, this conditionality may contribute to underestimating the impact of IB enrollment on students' probability of college enrollment. Given this assumption, if the probability of high school graduation was in fact equal for non-IB and IB students, the group of non-IB high school graduates would then include the more talented and motivated non-IB students who would have graduated anyway and the addition of the less talented and motivated students who would not have otherwise graduated. The subsequent lower probability of college enrollment among the weaker, less motivated non-IB students would result in a decrease in average college enrollment rates among non-IB students, leading in turn to an increase in the difference in the probability of college enrollment between IB and non-IB students.

CHALLENGE #4: DIFFERENTIAL PROBABILITY OF MISSING SEVENTH-GRADE ITBS SCORES

The fourth challenge to internal validity, like the second and third challenges, also arguably contributes to underestimation of the IB enrollment impact. IB and non-IB students do not have the same probability of having their seventh-grade ITBS mathematics and reading test score information. Of the original 20,422 students in the full sample, 20% of IB students as compared to 35% of non-IB students are missing their seventh-grade ITBS scores. Not including students in the comparison sample who are missing their ITBS information likely contributes to underestimating the IB enrollment impact. The logic is as follows. First, assume that students who are missing their seventh-grade ITBS scores have a higher probability of being from more mobile families than do

students with seventh-grade ITBS score information. A number of studies demonstrate that more mobile students have weaker academic performance (South, Haynie, & Bose, 2007; Stiefel, Schwartz, & Conger, 2010). If there is a greater probability that students missing their seventh-grade scores are mobile, and mobile students demonstrate weaker performance, then the inclusion in the comparison sample of students who are missing seventh-grade scores would result in “worse” average outcomes. Therefore, inclusion of students missing their seventh-grade test scores in the comparison sample would result in differences in average outcomes between IB and non-IB students that would likely be greater than the average differences in outcomes presented in Table 4.

When I fit models (1) and (2) for the full sample of CPS students attending high schools that offer the IB program (n=20,422), modifying the analytic strategy such that students’ ITBS scores are not included in the propensity-score selection model or as controls in the estimation model, the fitted estimates of the impact of IB enrollment are of substantially greater magnitude than are those shown in Table 4. This evidence supports the logic described in the paragraph above.¹⁷

CHALLENGE #5: SPILLOVER EFFECTS

The fifth internal validity challenge is the possibility of *spillover effects*; namely, the chance that the implementation of the IB program within the 13 CPS schools could affect students in the comparison sample who do not enroll in the IB program as defined in this study. For example, a teacher with IB training might use similar pedagogy and curriculum with both IB and non-IB students. This challenge would also lead to underestimation of the IB enrollment impact rather than overestimation, assuming that the teacher’s IB training might also have a positive impact on non-IB students’ outcomes.

CHALLENGE #6: EXTERNAL VALIDITY

The sixth validity challenge is a threat to external validity. Among IB-enrolled CPS students, 65% are eligible to receive free or reduced-price lunch and 74% are minorities. In comparison, as of 2010 approximately 20% of U.S. IB students are economically disadvantaged and 43% are minorities (IBO, 2010). The present results may only be relevant to students with demographic characteristics that are similar to those of students in this study’s sample and who attend high schools in settings similar to CPS.

DISCUSSION

ACADEMIC ACHIEVEMENT

In 2007, CPS district administrators established “ACT 20,” the goal for all 11th-grade CPS students to score at least a 20 on an unstandardized composite measure of their ACT subject-specific tests. They picked this score as it was considered generally to be the minimum necessary to matriculate at a four-year college in the state of Illinois (Easton, Ponisciak, Luppescu, & Consortium on Chicago School Research, 2008). ACT defines “College Readiness Benchmark Scores” for English, mathematics, reading, and science separately as 18, 22, 21, and 24 respectively. ACT predicts that students scoring at or above each of these benchmark scores will have a “50% or higher probability of earning a B or higher in the corresponding college course” (ACT, 2011).

Among non-IB students in the comparison sample, the mean (unstandardized) composite score on the ACT measure is 17.6 points with an SD of 3.3. This score is well below the Illinois state mean of 20.5 and the national average of 21.1 (for the 2007 cohort, cited in Easton et al., 2008). Based on this study’s propensity-score analysis—assuming, likely incorrectly—that Γ equals 1, I estimate that IB enrollment increases students’ ACT scores by 0.5 SDs—or 1.7 points (0.5×3.3)—to 19.1 points. Assuming an arguably more realistic level of Γ of 2, the lower bound of the 95% confidence interval around the IB enrollment impact is 0.08 SDs, indicating that when $\Gamma = 2$, IB enrollment increases students’ ACT scores by 0.3 points (0.08×3.3) to 17.9 points. These calculations suggest that while IB enrollment is contributing to CPS’s progress toward meeting the “ACT 20” goal, the practical impact is likely to be small.

HIGH SCHOOL GRADUATION

Improving high school graduation rates is one of the critical objectives of high school reform, nationally and in many school systems, including CPS (Allensworth & Easton, 2007). Rationales for this emphasis on high school graduation are based on studies showing that, compared to high school graduates, dropouts earn lower wages (Angrist & Krueger, 1991; Autor, Katz, & Kearney, 2005; Heckman, Lochner, & Todd, 2008), demonstrate a greater propensity to engage in criminal activities (Lochner & Moretti, 2004), and have a shorter life-expectancy (Lleras-Muney, 2005). Yet among the 26 nations in the Organization for Economic Cooperation and Development, as of 2004, the United States’s on-time high school graduation rate falls in 19th place (Goldin & Katz, 2008, p. 326). For the U.S. to remain internationally competitive, it will have to increase its graduation rates.

Even though ninth grade is the critical dropout period in CPS, I find that IB enrollment still has a meaningful impact on 11th-grade students' probability of graduation (Allensworth & Consortium on Chicago School Research, 2005). Even assuming extremely high levels of selection bias, this study estimates that IB enrollment substantially increases the probability that students will graduate from high school, contributing in a meaningful way to this critical objective of high school.

COLLEGE ENROLLMENT

The demands of today's global, technology-driven marketplace place a higher premium on workers with the sophisticated thinking and communicating skills that students gain through higher education (Goldin & Katz, 2008; Levy & Murnane, 2004). As Goldin and Katz (2008) demonstrate, from an individual economic perspective, a consequence of this "skill-biased technological change" is that as of 2005, each year of college education is associated with approximately a 14% lifetime rate of return. Given the increasing wage differential between high school graduates and college-educated workers, the U.S. must increase its supply of the latter if it is not to suffer from rising wage inequality. Other social benefits associated with some college, or college completion, include better health (Currie & Moretti, 2003) and greater civic participation (Dee, 2004; Verba, Schlozman, & Brady, 1995).

The IB program's reputation for successfully preparing students to enroll in college in large part substantiates its rapid growth. The results of this study indicate that IB deserves this reputation. Even when $\Gamma=2.5$, the p -value associated with the upper bound on the impact of IB enrollment on students' probability of college enrollment is less than 0.05. Therefore, though the estimate that IB enrollment increases students' probability of college enrollment by 38% is likely to be overstated, it is unlikely that selection bias could completely negate the impact of IB enrollment.

DIFFERENCES IN THE IB ENROLLMENT IMPACT BY GENDER

At the national level and in CPS, compared to boys, the probability that girls will graduate from high school is greater (Allensworth & Easton, 2007; Levin, Belfield, Muennig, & Rouse, 2007), as is the subsequent probability that girls will enroll in college (Charles & Luoh, 2003; Cho, 2007; Goldin & Katz, 2008). Jacob (2002) also shows that among low-income and minority students, compared to boys the probability that girls will enroll in college is 25% greater.

In the CPS context, I find that holding constant students' seventh-grade test percentiles, race, free-lunch eligibility, school, and cohort, the

odds that girls will enroll in the IB Diploma Program are twice the odds that boys will enroll. Yet, compared to IB-enrolled girls, the respective probabilities that male IB students will graduate from high school and enroll in college are 9.5 and 5.9 percentage points greater. One explanation for these findings is that only the most motivated and talented boys enroll in the IB program, while there is a wider distribution of motivation and talent among IB-enrolled girls. The data, however, show that among IB students, boys' scores on the ITBS mathematics test are only 3 percentiles higher and are not statistically distinguishable from those of girls on the ITBS reading test. This evidence suggests that differences in the impact of IB enrollment by gender are not driven purely by compositional differences between boys and girls.

ADVANCED PLACEMENT

An evaluation of the IB Diploma Program is remiss to not reference the College Board's 34 Advanced Placement (AP) courses and examinations. Like IB Diploma Program courses, AP courses have a reputation for rigorous high school level preparation for college, and students typically enroll in them during the 11th and 12th grades.

However, a number of factors differentiate the two options. First, AP has a much larger U.S. student audience than does the IB Diploma Program. Among U.S. high school students, in 2011 nearly 30% took at least one AP examination (College Board, 2013), compared to far lower proportions of students who take at least one IB examination. Second, whereas AP can play a small role in a student's high school experience, enrollment in the IB Diploma Program is all consuming. AP students often take just one or two AP courses and examinations. In contrast, to be eligible to earn the IB Diploma, IB students must take six IB subject-matter courses and the TOK epistemology course, write the 4,000-word Extended Essay, and participate in the Creativity, Action and Service program. Third, unlike IB courses, which follow a specified curriculum and pedagogy, AP courses do not follow a standardized curriculum and there is great variation in the approaches and texts that AP teachers use to deliver their course content.

Jackson (2010) uses causal methods to estimate that AP course enrollment increases the probability of scoring above 24 on the ACT by a significant 13%. He also finds that compared to students who do not enroll in AP courses, the probability that students who enroll in AP courses will subsequently enroll in college is a marginally significant 5% greater. He presents suggestive evidence that the relative improvements are greatest for minority students. In the CPS setting, however, Roderick et al. (forthcoming) use

qualitative and descriptive quantitative methods to argue that the impact of AP coursework on college outcomes has been very limited in its effects on college attendance, with no effects on college persistence.

IMPLICATIONS FOR FUTURE RESEARCH AND POLICY

The results suggest the value of further study of the mechanisms that drive the overall IB enrollment impact and the differences in the IB impact by gender. Mechanisms driving the overall impact could include: IB professional development, pedagogy, and curriculum; peer effects; and setting high expectations for students. Mechanisms driving the gender differential could include: differences in course-taking patterns by gender; teachers' setting higher expectations for boys than for girls; and/or stronger positive peer effects among boys than among girls. Greater understanding of how the IB enrollment impact works would suggest to policy makers and other educational stakeholders ways to provide broader student audiences with the opportunities for educational access afforded by IB enrollment. Better understanding of how enrollment in the IB program successfully increases boys' educational attainment would promote greater understanding of mechanisms that could attenuate the gender gap in educational attainment.

COST OF THE IB DIPLOMA PROGRAM EFFECTS

The IB Diploma Program is one of several educational strategies designed to improve students' secondary academic success, each of which have temporal and monetary costs. For example, Jackson (2010) estimates that an AP program costs schools approximately \$200 per student. An important policy consideration is whether the costs for implementing the IB program justify the benefits of IB enrollment for students.

According to the current CPS IB Diploma Program administrator (personal correspondence, December 14, 2010), as of 2010, a year of CPS's implementation of the official 11th- through 12th-grade IB Diploma Program—and associated preparatory programs for students in the ninth and 10th grades—costs approximately \$650,000. In 2010, each of the 13 high schools included in this study received \$10,000, which covered their 2011 IBO annual fee. CPS invested an additional \$200,000 in supplemental fees, which were distributed across the 13 high schools and used to cover costs such as textbooks and IB-sponsored teacher training. Fees to register for the examinations and subject-specific examination fees cost approximately \$320,000.

Of the \$650,000, CPS pays \$490,000 and the Illinois Department of Education reimburses the district for the balance.¹⁸ The \$650,000 funds

the IB Diploma Program and related preparation courses for the approximately 3,000 students who took one or more IB Diploma Program and/or associated preparatory courses, and covers all participating 12th graders' examination registration and subject-specific test fees.¹⁹ Therefore, in 2010, the average cost-per-student for implementing the IB Diploma Program and related preparation courses in CPS high schools, and for permitting eligible students to take IB Diploma examinations, was approximately \$217 ($\$650,000/3,000$). This estimate should be understood cautiously, as it is possible that the \$650,000 might underestimate costs such as the time required for teachers to participate in IB professional development training and plan their courses as well as miscellaneous costs for space, materials, etc.

The Economic Policy Institute estimates that on average, workers without a high school diploma make approximately \$22,000 per year and high school graduates with no college education, on average, make approximately \$29,000 per year (Mishel, Bernstein, & Shierholz, 2009). Therefore, assuming a 40-year work span and a 3% discount rate, the net present value of the lifetime earnings gap is approximately \$162,000. I find that IB enrollment increases the probability of high school graduation by approximately 21% and that this result is extremely robust to high degrees of selection bias. Therefore the results suggest that the net present value of lifetime earnings induced by IB Diploma Program enrollment is approximately \$34,020 ($\$162,000 * 0.21$), which amounts to a benefit-cost ratio of approximately 78 to 1 [$\$34,020 / (\$217 * \text{two years of IB enrollment})$].

This calculation can be understood as an underestimate of the benefits of investment in the IB Diploma Program, as it does not take into account additional individual lifetime earning benefits caused by the effect of enrollment in the IB Diploma Program on college enrollment—people with “some college” on average, make approximately \$33,000 per year (Mishel et al., 2009). Nor does the calculation take into account the social gains of increasing high school graduation rates, including increased tax payments, and savings in public health and prison expenditures (Levin et al., 2007). However, as mentioned, the \$650,000 annual cost figure might underestimate the real cost of running the program, in which case the returns could be somewhat lower.

The present study does not permit comparison of IB results to Jackson's (2010) AP results because the way we define the IB versus AP treatments is not consistent, and because with the sensitivity analysis, it is not possible to estimate a specific magnitude of the IB enrollment impact. As such, it is not possible to compare AP versus IB cost-effectiveness using Jackson's estimated cost of \$200 per student to implement the AP

program compared to this study's similar calculated cost-per-student to implement the IB Diploma Program. It is worth noting that the IB costs include curricular materials for teachers and a holistic experience for students, which is not the case for AP.

CONCLUSION

This study shows that IB enrollment increases students' academic achievement, probability of high school graduation, and probability of college enrollment. Though selection bias may contribute to overstating the estimates, the conclusion from the sensitivity analyses is that it is unlikely that this internal-validity challenge negates the principal finding.

The results are valuable for three reasons. First, the IB program has grown rapidly in U.S. public schools over the past 15 years, and expected continuation of this trajectory is based on widespread perceptions of the program's success as a college preparation intervention. Therefore, understanding that the IB Diploma Program is indeed helping CPS to accomplish the primary purposes of high school—to increase student's academic achievement during high school and to encourage them to graduate from high school and enroll in college—provides policy makers and educators, as well as students and parents, with valuable information with which to make decisions about future investments in IB. Second, the results contribute to knowledge of the means through which to improve high school education for disadvantaged urban youth. Given that CPS is a large urban setting, attended primarily by minority and low-income students, the findings suggest that the IB Diploma Program may be a means through which to increase students' high school achievement, probability of high school graduation, and probability of college enrollment in similar settings. Finally, the results suggest that IB enrollment is especially beneficial for boys, for whom the probability of graduating from high school and enrolling in college—in CPS and at the national level—is substantially less than for girls.

Acknowledgments

I am grateful to Sara Leven from Chicago Public Schools for her comprehensive support. I am thankful to Amy Novell from Chicago Public Schools and to Sue Spote and Vanessa Coca from the Consortium on Chicago School Research for generously providing me with access to the data I use in this project. I especially thank Meira Levinson, Richard Murnane, John Willett, and Juan Esteban Saavedra for their advice, time, and support. Thank you as well to the anonymous reviewers that helped me to make this study suitable for publication. I thankfully acknowledge

financial support from the Center for the Advancement and Study of International Education, the Frederick Sheldon Harvard Travel Fellowship, the Fulbright Program, and the Harvard Graduate School of Education. Any errors are mine.

Notes

1. Abbreviations: International Baccalaureate (IB); International Baccalaureate Organization (IBO); Chicago Public Schools (CPS). Note that ACT is not an abbreviation; it is the full name of the college entrance examination.

2. The ACT organization administers the ACT college admissions examination, which includes English, mathematics, reading, and science sections (ACT, 2011).

3. www.ibo.org, retrieved on January 23, 2013. Indianapolis currently has six high school level IB programs, followed by five each in Philadelphia, Minneapolis, and Saint Paul. Two of Chicago's 14 currently operating IB Diploma Programs opened in 2012. Data from those two schools are not available to include in the analyses presented in this paper. During the timeframe of the data included in this paper, one IB-offering school was open that has since closed.

4. See Heckman and LaFontaine (2010) for a discussion of U.S. high school graduation rates.

5. In the 2005-2006 school year, the Illinois Standard Achievement Tests replaced the ITBS for seventh-grade testing (Easton et al., 2003).

6. The majority of CPS students who enroll in college attend postsecondary institutions located in Illinois (Roderick, 2006). According to the NSC's Assistant Director of Research Services, "the NSC post-secondary enrollment data coverage for students in Illinois is higher than the national average of 92%" (personal correspondence, February 23, 2011). By cross-referencing NSC data with CPS exit-survey data, Roderick, Nagaoka, Allensworth, Stoker, Correa, and Coca (2006) estimate that the NSC data "misses approximately 5 percent of CPS graduates who may be enrolled in college" (cited in Roderick, 2006, footnote 56). Roderick et al. also characterize the Illinois colleges for which the NSC does not verify enrollment as "primarily local proprietary and technical institutions" (2006). Given the assumption that compared to IB students, the probability is greater that non-IB students will attend local proprietary and technical institutions, missing NSC enrollment data could contribute to a slight understatement of the percentage of non-IB students who enroll in college.

7. This study does not have access to records for the 2005-2006 11th-grade cohort. There is no reason to believe that the results of the present analysis would differ if the 2005-2006 records were included in the full sample.

8. The data used in this study includes ACT subject scores, but not the overall ACT Composite score, so the analysis required creation of the Composite score. Under the website heading, "How ACT figures the multiple-choice test scores and the Composite score," ACT explains to a student audience that, "The Composite score is the average of your four test scores, rounded to the nearest whole number" (ACT, 2013).

9. The requirements for the CPS high school degree are more demanding than the requirements for an Illinois degree. The data do not indicate whether students later earned the General Education Development equivalency degree (GED).

10. Though these proportions seem high, these students persisted beyond ninth grade, the year in which CPS students are at the highest risk of dropout (Allensworth, 2005).

11. I do not qualify my definition of the IB treatment by whether students enrolled in official IB preparatory programs in ninth and/or 10th grade because the official IB Diploma Program begins in 11th grade and because of data limitations.

12. The selection model includes quadratic functional forms of the mathematics and reading percentiles because this more flexible form ought to be a better fit for IB students' ITBS scores. Estimates of the IB enrollment impact are statistically indistinguishable in models that include respective nested linear, quadratic, cubic, and quartic functional forms of the ITBS percentiles.

13. In practice, each observation is weighted by the square of the weights (Imbens & Wooldridge, 2009).

14. It is unlikely that the types of unmeasured characteristics that affect students' self-selection into IB will contribute to underestimation of the IB enrollment impact.

15. For a technical explanation of how the Wilcoxon and McNemar tests constitute the foundation of this sensitivity analysis, see Rosenbaum (2002).

16. I compare my sensitivity analysis results to the results of fitting a maximum likelihood Heckman's correction model. If the model is correctly specified, the fitted Heckman estimates are free of selection bias from observables and unobservables. The model in which probability of high school graduation is the outcome predicts a small and insignificant rho of 0.008, which reinforces my high school graduation finding. When ACT scores and probability of college enrollment are the outcomes, the rhos are larger and significant, potentially suggesting that selection drives the impact on students' ACT scores and probability of college enrollment to a great extent. However, the selection model probably is not correctly specified. Further, in comparison to the more flexible fit afforded by the non-parametric propensity score approach, the (linear) Heckman model may be a worse fit because it is subject to strict normality assumptions.

17. The result of fitting Models (1) and (2) for the full sample of 20,422 students is to find that IB enrollment increases students' probability of taking the ACT test by 13 percentage points, their ACT test scores by 1.4 SDs, their probability of high school graduation by 21 percentage points, and their probability of college enrollment by 33 percentage points.

18. The Illinois Department of Education reimburses 90% of the subject-specific examination fees.

19. According to the CPS IB Diploma Program administrator (personal correspondence, December 14, 2010), in 2010, 935 of these 3,000 students enrolled in the full IB Diploma Program, as indicated by their enrollment in TOK. Therefore, the 2,065 students who only partially enrolled in the IB Diploma Program and preparatory programs essentially subsidize the costs of the 935 students who enroll in the full IB Diploma Program and take the IB examinations.

References

- ACT. (2011). *College Readiness Standards for EXPLORE, PLAN and the ACT*. Retrieved on January 29, 2013, from: <http://www.act.org/standard/>
- ACT. (2013). Retrieved on January 29, 2013, from: <http://www.act.org/>
- Allensworth, E. (2005). *Graduation and dropout trends in Chicago: A look at cohorts of students from 1991 through 2004*. Chicago, IL: Consortium on Chicago School Research.
- Allensworth, E., & Easton, J. (2007). *What matters for staying on track and graduating in Chicago public high schools: A close look at course grades, failures and attendance in the freshman year*. Chicago, IL: Consortium on Chicago School Research.
- Allensworth, E., & Rosenkranz, T. (2000). *Access to Magnet Schools in Chicago*. Chicago, IL: Consortium on Chicago School Research.
- Angrist, J. & Krueger, A. (1991). Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics*, 106(4), 979–1014.
- Autor, D. H., Katz, L. F., & Kearney, M. S. (2005). *Rising wage inequality: The role of composition and prices* (NBER technical working paper No. 11627).
- Burris, C. C., Welner, K. G., Wiley, E. W., & Murphy, J. (2007). A world-class curriculum for all. *Educational Leadership*, 64(7), 53–56.
- Byrd, S., Ellington, L., Gross, P., Jago, C., & Stern, S. (2007). *Advanced Placement and International Baccalaureate: Do they deserve gold stars?* Washington, DC: Thomas B. Fordham Institute.
- Cech, S. J. (2008). World grows smaller, IB gets BIG. *State Legislatures*, 34(2), 20–24.
- Charles, K. K., & Luoh, M. (2003). Gender differences in completed schooling. *The Review of Economics and Statistics*, 85(3), 559–577.
- Chicago (IL) Board of Education. (1867). *Official report of the proceedings of the Board of Education of the city of Chicago*. Chicago, IL: Chicago Public Schools.
- Chicago School Reform Board of Trustees. (1999). *Chicago Public Schools fiscal year 2000 final budget: Adopted July 23, 1999*. Chicago, IL: Chicago Public Schools.
- Cho, D. (2007). The role of high school performance in explaining women's rising college enrollment. *Economics of Education Review*, 26(4), 450–462.
- Coca, V., Kelley-Kemple, T., Roderick, M., & Nagaoka, J. (2011). *Chicago Public School high school students in IB programs and their transition to college*. Chicago, IL: Consortium on Chicago School Research.
- College Board. (2013). *AP Program Facts*. Retrieved on January 29, 2013, from: <http://press.collegeboard.org/ap/fact-sheet>
- Currie, J., & Moretti, M. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *Quarterly Journal of Economics*, 118(4), 1495–1532.
- Dee, T. (2004). Are there civic returns to education? *Journal of Public Economics*, 88(9), 1697–1720.
- Easton, J. Q., Correa, M., Lupescu, S., Park, Ponisciak, S., H. S., Rosenkranz, T., & Sporte, S. (2003). *How do they compare? ITBS and ISAT reading and mathematics in Chicago Public Schools, 1999-2002*. Chicago, IL: Consortium on Chicago School Research.
- Easton, J. Q., Ponisciak, S., & Lupescu, S. (2008). *From high school to the future: The pathway to 20*. Chicago, IL: Consortium on Chicago School Research.
- Foust, R. C., Hertberg-Davis, H., & Callahan, C. M. (2009). Students' perceptions of the non-academic advantages and disadvantages of participation in Advanced Placement courses and International Baccalaureate programs. *Adolescence*, 44(174), 289–312.
- Goldin, C., & Katz, L. (2008). *The race between education and technology*. Cambridge, MA: The Belknap Press of Harvard University Press.

- Heckman, J. J., & LaFontaine, P. A. (2010). The American high school graduation rate: Trends and levels. *Review of Economics and Statistics*, 92(2), 244–262.
- Heckman, J. J., Lochner, L. J., & Todd, P. E. (2008). Earnings Functions and Rates of Return. *Journal of Human Capital*, 2(1), 1-31.
- Hirano, K., Imbens, G. W., & Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4), 1161–1189.
- Illinois Department of Education. (2004). *Illinois ACT Scores Rise: Second straight year since all public high school juniors began taking ACT*. Retrieved on February 24, 2011, from: www.isbe.state.il.us/news/2004/aug18_04.htm
- Illinois State Board of Education. (2009). *New U.S. Department of Education Race and Ethnicity Data Standards* (Powerpoint presentation). Retrieved on February 4, 2013, from: http://www.isbe.state.il.us/research/Default.htm
- Illinois State Board of Education (February 5, 2013). Informal interview.
- Imbens, G., & Wooldridge, J. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1), 5–86.
- International Baccalaureate Organization. (2010). Unpublished IB enrollment statistics.
- International Baccalaureate Organization. (2012). *IB Annual Review*. Retrieved from http://ibo.org/facts/annualreview/
- International Baccalaureate Organization. (2013). Retrieved on February 5, 2013, from: http://www.ibo.org/
- Jackson, C. K. (2010). A little now for a lot later. *Journal of Human Resources*, 45(3), 591–639.
- Jacob B. A. (2002). Where the boys aren't: Non-cognitive skills, returns to school and the gender gap in higher education. *Economics of Education Review*, 21(6), 589–598.
- Keele, L. (2009). *An overview of r-bounds: An R package for Rosenbaum bounds sensitivity analysis with matched data*. Unpublished manuscript.
- Kieran, J. (2007). Factors influencing students' choice(s) of experimental science subjects within the International Baccalaureate Diploma Programme. *Journal of Research in International Education*, 6(9), 9–39.
- Kyburg, R. M., Hertberg-Davis, H., & Callahan, C. M. (2007). Advanced Placement and International Baccalaureate programs: Optimal learning environments for talented minorities? *Journal of Advanced Academics*, 18(2), 172–215.
- Levin, H. M., Belfield, C., Muennig, P., & Rouse, C. (2007). The public returns to public educational investments in African-American males. *Economics of Education Review*, 26(6), 699–708.
- Levy, F., & Murnane, R. J. (2004). *The new division of labor: How computers are creating the next job market*. Princeton, NJ: Princeton University Press.
- Lleras-Muney, A. (2005). The relationship between education and adult mortality in the United States. *Review of Economic Studies*, 72(1), 189–221.
- Lochner, L., & Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American Economic Review*, 94(1), 155–189.
- Marburger, D. (2006). Does mandatory attendance improve student performance? *Journal of Economic Education*, 37(2), 148–155.
- Mathews, J., & Hill, I. (2005). *Supertest: How the International Baccalaureate can strengthen our schools*. Chicago, IL: Open Court.
- McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2), 153–157.
- Mishel, L. R., Bernstein, J., & Shierholz, H. (2009). *The state of working America: 2008-2009*. New York, NY: Economic Policy Institute.
- Murnane, R., & Willett, J. (2011). *Methods matter: Improving causal inference in educational and social science research*. New York, NY: Oxford University Press.

- National Student Clearinghouse. (2011). Retrieved on February 23, 2011, from: <http://www.studentclearinghouse.org/>
- Rice, W. K. (1995). *The state report cards for 1995-1995: An analysis of the first decade*. Data compliance office of accountability. Chicago, IL: Chicago Public Schools.
- Riverside Publishing. (2010). *Iowa Test of Basic Skills Forms A, B and C: Product Details*. Retrieved on February 5, 2013, from: <http://www.riversidepublishing.com/products/itbs/details.html>
- Roderick, M. (2006). *Closing the aspirations-attainment gap: Implications for high school reform, a commentary from Chicago*. MDRC.
- Roderick, M., Coca, V., Moeller, E., Kelley-Kemple, T., & Nagaoka, J. (forthcoming). *Senior year*. Chicago, IL: Consortium on Chicago School Research.
- Roderick, M., Nagaoka, J., Allensworth, E., Stoker, G., Correa, M., & Coca, V. (2006). *From high school to the future: A first look at Chicago Public School graduates: College enrollment, college preparation, and graduation from four-year colleges*. Chicago, IL: Consortium on Chicago School Research.
- Roderick, M., Nagaoka, J., Coca, V., & Moeller, E. (2009). *From high school to the future: Making hard work pay off: The road for students in academically advanced programs*. Chicago, IL: Consortium on Chicago School Research.
- Rosenbaum, P. R. (2002). *Observational studies* (2nd ed.). New York, NY: Springer.
- Siskin, L. S., & Weinstein, M. (2008). *Supplemental survey to: Creating support structures and services for Title I high schools implementing the International Baccalaureate programs*. New York University, New York, NY: Institute for Education and Social Policy.
- South, S. J., Haynie, D. L., & Bose, S. (2007). Student mobility and school dropout. *Social Science Research*, 36(1), 68–94.
- Spittle, B., Leven, S., & Roderick, M. (2008). IB as a college access strategy in Chicago Public Schools. *International Baccalaureate Regional Conference*. San Francisco, CA.
- Stables, A., & Sian, S. (1995). Gender differences in students' approaches to A-level subject choices and perceptions of A-level subjects: A study of first-year A-level students in a tertiary college. *Educational Research*, 37(1), 39–51.
- Stiefel, L., Schwartz, A. E., & Conger, D. (2010). Age of entry and the high school performance of immigrant youth. *Journal of Urban Economics*, 67(3), 303–314.
- U.S. Department of Agriculture. (2012). *Eligibility Manual for School Meals: Determining and Verifying Eligibility*. Washington, DC: Child Nutrition Programs, Food and Nutrition Service. Retrieved on February 5, 2013, from: <http://www.fns.usda.gov/cnd/lunch/>
- U.S. Department of Education. (2007). *The nation's report card: America's high school graduates: Results from the 2005 NAEP high school transcript study*. Washington, DC: Government Printing Office.
- U.S. Department of Education. (2011a). *A blueprint for reform: The reauthorization of the elementary and secondary education act*. Alexandria, VA: EdPubs. Retrieved on January 17, 2011, from: <http://www2.ed.gov/policy/elsec/leg/blueprint/blueprint.pdf>
- U.S. Department of Education. (2011b). *Race to the top fund: States' applications for phase 2*. Retrieved on February 24, 2011, from: <http://www2.ed.gov/programs/racetothetop/phase2-applications/index.html>
- Verba, S., Schlozman, K. L., and Brady, H. E. (1995). *Voice and equality: Civic voluntarism in American politics*. Cambridge, MA: Harvard University Press.

APPENDIX

Parameter estimates (standard errors), sample sizes, and approximate p -values for 20 separate WLS-fitted models that describe the impact of IB enrollment on four student outcomes across cohorts and by cohort.

	Outcome Variables			
	(1)	(2)	(3)	(4)
	Takes the ACT Examination	Academic Achievement (ACT Scores)	Probability of High school Graduation	Probability of College Enrollment
All cohorts	0.034*** (0.011) n=11,592	0.536*** (0.051) n=11,148	0.173*** (0.018) n=11,592	0.217*** (0.024) n=11,592
2003	0.059*** (0.015) n=2,592	0.527*** (0.088) n=2,432	0.181*** (0.017) n=2,595	0.239*** (0.033) n=2,595
2004	0.042** (0.015) n=2,868	0.549*** (0.070) n=2,761	0.179*** (0.025) n=2,868	0.182*** (0.041) n=2,868
2005	0.029** (0.011) n=2,906	0.587*** (0.074) n=2,815	0.168*** (0.029) n=2,906	0.287*** (0.047) n=2,906
2007	0.022 (0.016) n=3,223	0.515*** (0.057) n=3,140	0.173*** (0.025) n=3,223	0.173*** (0.025) n=3,223

Note: All terms included in Model (2) are included in the fitted models.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

ANNA ROSEFSKY SAAVEDRA is an Associate Policy Researcher at the RAND Corporation. Her research interests include education policy, educational equity, educational program evaluations, civic education, 21st-century skills education, and international education. Recent publications include “Do Colleges Cultivate Critical Thinking, Problem Solving, Writing and Interpersonal Skills?”, co-authored with Juan Esteban Saavedra, in *Economics of Education Review* (2011), and “From Dry to Dynamic Civics Education Curricula” in *Making Civics Count* (2012), published by Harvard Education Press.